DOCUMENT RESUME

ED 307 593 CS 009 676

AUTHOR Stallman, Anne C.; And Others
TITLE Are "New" Words Really New? Technical Report No. 471.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE May 89
GRANT OEG-0087-C1001
NOTE 24p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Analysis of Variance; *Basal Reading; Elementary Education; Grade 2; Grade 5; Reading Comprehension; Reading Research; *Vocabulary Skills

ABSTRACT

A study examined whether students already know the meanings of reading vocabulary presented as "new" vocabulary in their basal readers. Subjects, 142 second and fifth grade students from two midwestern school districts, were asked to complete a yes/no test and a multiple choice test of vocabulary in their basal readers at their grade level and at one and two years beyond their grade levels. Results indicated that children in both groups did not perform much better on the words they had been formally taught than on the words they had not yet been taught. Findings suggest that educators and publishers should reevaluate the criteria they use to select vocabulary for instruction in basal readers. (Seven tables of data are included, and 25 references are attached.) (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED307598

CENTER FOR THE STUDY OF READING

Technical Report No. 471

ARE "NEW" WORDS REALLY NEW?

**Anne C. Stallman, Michelle Commeyras,
Bonnie M. Kerr, Kathryn Meyer-Reimer,
Robert Jimenez, Douglas K. Hartman, and
P. David Pearson
University of Illinois at Urbana-Champaign**

May 1989

**University of Illinois at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820**

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

P. ANDERSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The work upon which this publication was based was supported in part by the Office of Educational Research and Improvement under Cooperative Agreement No. OEG 0087-C1001, with the Reading Research and Education Center. The publication does not necessarily reflect the views of the agency supporting the research.

CS009676

BEST COPY AVAILABLE

EDITORIAL ADVISORY BOARD
1988-89

Beck, Diana

Commeyras, Michelle

Foertsch, Daniel

Hartman, Doug

Jacobson, Michael

Jehng, Jihn-Chang

Jimenez, Robert

Kerr, Bonnie

Kerr, Paul

Meyer, Jennifer

Moran, Juan

Ohtsuka, Keisuke

Roe, Mary

Schommer, Marlene

Scott, Judy

Stallman, Anne

Wilkinson, Ian

Wolff, Phillip

MANAGING EDITOR
Mary A. Foertsch

MANUSCRIPT PRODUCTION ASSISTANTS
Delores Plowman
Nancy Diedrich

Abstract

The present study investigates the question of whether second and fifth grade children already know the meanings of reading vocabulary presented as "new" vocabulary in their basal readers. Students from two midwestern school districts (N=142) were asked to complete a yes/no test and a multiple choice test to assess their knowledge of vocabulary presented in their current grade level basal readers as well as their knowledge of vocabulary presented in basal readers 1 and 2 years beyond their grade level. Overall mean scores for both second and fifth grade cohorts were above 75% for grade level words and over 70% for the "new" words which had not yet been presented. Children were found not to perform much better on those words formally taught in basal reading series than on words they had not yet been taught. It is suggested that educators and publishers need to reevaluate the criteria they use to select vocabulary for instruction in basal readers.

ARE "NEW" WORDS REALLY NEW?

The teaching of vocabulary words *before* reading a selection is as sacred a canon of reading instruction as is asking questions *after* reading the selection. The rituals of pre-teaching vocabulary words and holding post-reading discussion sessions are regularly practiced in the name of reading instruction. In short, conventional classroom wisdom has upheld these two practices as important in learning to read. It was not until recent times, however, that vocabulary knowledge has been shown to have a critical role in reading comprehension. Repeatedly, scholars (e.g., Adams & Huggins, 1985; Anderson & Freebody, 1982; Beck & McKeown, in press; Johnson & Pearson, 1984; Nagy, Herman, & Anderson, 1985) have demonstrated the strength of this relationship: Knowledge of vocabulary words helps us understand when we read. Based upon current thought and the collective wisdom and intuitions of teachers and readers over the years, "new" (or "unknown") words are taught to students with the belief that it will help them better understand what they read.

The assumption underlying the pre-teaching of vocabulary to facilitate comprehension is that the pre-taught words are "new" to students. But what if this assumption goes unexamined and known words are taught as "new" words to students? Such an oversight would mean that teachers are spending needless time teaching students what they already know. In thinking about where such a problem might occur, commercially developed reading programs seemed a likely place to start. In the teachers' manuals of these programs, the authors provide lists of "new" words and suggestions for pre-teaching them to students. We began by wondering if it were possible that both developers and users of reading lessons have assumed that the words designated as "new" in the basal series are unknown to students when, in fact, the words may not be new. "New" words are generally defined as those words that have not appeared in previous readers. But does it follow that because words have not appeared in previous readers, students do not know what they mean? As a result, this study was designed to investigate whether children already knew many of the "new" words that were recommended for instruction in teacher's manuals. Toward the end of the school year, after students had finished most of the reader they were working on, we compared their knowledge of "new" vocabulary from their present basal reader (words that they were supposed to have learned) with their knowledge of the "new" words from the readers for the following two grades (words that they were not expected to have learned).

A study like this has historical precedents. Over a quarter century ago, Gates (1961, 1962) demonstrated that while second grade students knew over 80% of the second grade words for which they had received instruction, they also knew over 75% of the third and fourth grade words for which they had not received instruction. He concluded that much of the vocabulary recommended for teaching in basal manuals was already known by the students. One could easily conclude the same is true today, but there is evidence to suggest that vocabulary in basals has changed since Gates' research in the late fifties. Recent research suggests that the number of vocabulary words presented in current basal series has increased. Morris and Johns (1987) found that there had been a sizable increase from the 1960s to the 1980s in the vocabulary load in first grade materials for the three reading series (Ginn and Company, Houghton Mifflin, and Scott Foresman) they examined. They also found, possibly as a result of the increased vocabulary load, that readability levels were somewhat higher in the 1980s than they had been during the 60s. If vocabulary in basal series is more difficult, then it is possible that today's students will not know as many of the "new" words as did their counterparts 25 years ago.

Perhaps the basal recommendations to teach students what they already know would not warrant concern if teachers did not pay attention to them. However, Durkin (1984) found that teachers are likely to use the section of their manuals on vocabulary instruction which recommends the introduction of "new" vocabulary words. The irony of Durkin's finding is that while researchers have shown that children's vocabulary knowledge *can* be increased through some non-traditional instructional techniques emphasizing conceptually-driven activities (Duffelmeyer, 1984; Gipe, 1978-79; McKeown, 1985; Nagy, Herman, & Anderson, 1985; Searls & Klesius, 1984), the instruction found in basals is much more conventional, with a clear emphasis on definitional approaches.

Researchers interested in the area of vocabulary have employed a number of different ways to assess word knowledge. Multiple choice tests are most commonly used to assess vocabulary knowledge. There are several ways to design multiple choice tests which are sensitive to partial word knowledge (Kolstad, Briggs, Bryant, & Kolstad, 1983; Nagy, Herman, & Anderson, 1985). Measuring partial knowledge is important because research has shown that acquisition of vocabulary knowledge is incremental in nature (Nagy, Herman, & Anderson, 1985). One way to measure partial word knowledge is to employ a format which allows the students to select more than one right answer (Nimmer, 1983). A format with multiple correct responses makes it possible to give a child credit for knowing some aspects of the word even if she does not have a complete adult-like understanding of it.

Another way to tap partial knowledge is to use different types of questions. Items that include context, or use examples, or attributes of the word are effective ways of measuring partial as well as complete understanding of word meanings (Anderson & Kulhavy, 1972; Duffelmeyer, 1984; Nagy, Herman, & Anderson, 1985).

Anderson and Freebody (1982, 1985) have found the yes/no test, in which the student indicates that she does or does not know the meaning of a particular word, to be an efficient and accurate means of determining the percentage of words in a particular list for which a child has some knowledge. Although the yes/no test is sensitive to partial knowledge of word meanings, it does not provide a way of ascertaining *which* meanings of particular words the child knows.

Our study was designed to replicate and extend the vocabulary studies conducted by Gates in the late 1950s. We set out to answer the same question Gates investigated: Do children know many of the "new" words in basal readers? In order to establish comparability with Gates work, we included second grade students in our sample.

Our study extends Gates' work in several ways. First, we used multiple measures of vocabulary knowledge. Also, we looked at more than one basal series; we included words from the 1958 Macmillan Series that Gates had used as well as recent editions of the Scott Foresman and Houghton Mifflin reading series that were used in the students' reading programs. Gates tested second and third graders: we tested second graders and fifth graders in our study because we thought it would be informative to examine the possibility that the trends for primary age children might be different than those for intermediate age children.

Method

Subjects

Students from two midwestern school districts participated in this study. In District A, students were selected from two second grade classrooms (N=45) and two fifth grade classrooms (N=40). In District B, one second (N=28) and one fifth grade (N=29) class participated. Students in all six classrooms were grouped homogeneously for reading instruction. The teachers rated the majority of the students as average readers. District A had a 42% minority student population while District B had a minority student population of 20%. None of the students in this study had been identified by their schools as having educational disabilities.

Instructional Materials

The participating classrooms used the basal reading series adopted by the entire school district. District A used the Scott Foresman and Company Series (Aaron, Jackson, Riggs, Smith, & Tierney, 1983); District B used the Houghton Mifflin Company Series (Durr, LePere, Pikulski, & Bunyan, 1983). Districts with different reading series adoptions were chosen in order to evaluate the consistency of findings across reading series. This study was conducted near the end of the school year.

At that time the second-grade students were finishing the 2_2 level readers, and the fifth-grade students were completing the fifth-grade books.

Testing

Two dependent measures were used to assess the students' knowledge of "new" vocabulary words from their basal reading series. The primary measure was a multiple choice test with more than one right answer, designed to assess students' knowledge of specific meanings of words. The multiple choice items tested the same meaning that the teacher's manual designated for instruction.

The multiple choice tests were comprised of words from the following categories:

1. twenty words from the students' current basal instructional level;
2. twenty words from one grade level above the students' current basal instructional level; and,
3. twenty words from two grade levels above.

These 60 words were randomly selected from the lists of "new" vocabulary words in the teacher's manuals. Each word was randomly assigned to one of four question types: straight definition, definition with context, examples, or attributions of the word (See Figure 1 for examples).

Each test item had five options with either one or two correct answers. Each distractor for the vocabulary items met one of these criteria:

1. the same semantic category as the target word;
2. the same syntactic category as the target word;
3. graphically similar to a decoding miscue of the target word;
4. unrelated to the target word.

[Insert Table 1 about here.]

The second dependent measure was a 180 item yes/no test (Anderson & Freebody, 1981, 1983, 1985). In the yes/no vocabulary test, subjects were instructed to circle each word for which they knew the meaning. The yes/no tests were comprised of words from the following three categories:

1. the same 60 words that were on multiple choice tests;
2. thirty words from the Macmillan (1958) Series that Gates used in his studies (10 from the level at which students were reading and 10 from each of the next two grade levels);
3. ninety pseudowords.

Pseudowords were constructed in accordance with the Anderson and Freebody method (Anderson & Freebody, 1981). For some of the pseudowords, vowels in real words were changed (e.g., blink became blonk); for others, permissible, but not actual, English letter strings were constructed (grack). A third category of pseudowords was constructed by adding inappropriate affixes to real words (departness). The pseudo- and real words were matched for number of syllables and letters.

Four forms of each multiple choice test and yes/no test were used to minimize the possibility of successful copying and to avoid effects due to the order of the test items. Test form had no effect on student scores; hence it was dropped from further analyses.

Scoring

Two scoring systems were used to score the multiple choice tests of vocabulary knowledge. One system used the logic of a true/false algorithm that scored every option. Students were given credit for each keyed response that they selected. They were also given credit for every incorrect response that they did not select. This system, in effect, transformed the 60-item multiple choice test into a 300-item true/false test because all five options for each item were scored. Due to the variable number of correct answers (one or two) this was an appropriate method for scoring the students' responses.

The second scoring system, designed to mimic a more conventional multiple choice test, gave credit for an item if the student picked either one or two correct answers and did not choose any incorrect answers. Correlations between the two scoring systems revealed that they yielded essentially the same information about the students' vocabulary knowledge ($r = .94$; $p < .001$).

The yes/no tests were originally scored using the conventional Anderson/Freebody formula:

$$\frac{P(H) = P(FA)}{1 - P(FA)}$$

Using this formula, a student's score is calculated as follows. The proportion of pseudowords that the student marks as known [P(FA)] is subtracted from the proportion of real words [P(H)] that the student marks as known. The resulting number is then divided by 1 minus the proportion of nonwords that the student marks as known [P(FA)]. This gives an estimate of the proportion of words tested that the student truly knows, after correcting for guessing.

However, the Anderson/Freebody formula poses a problem when a student circles all the real words as known. In this case, a student would receive a score of 100 % irrespective of the number of pseudo words that were also circled as known (see Table 2). This was a concern in this study because there were students who circled all the real words as known. The following formula, which we used to score the yes/no test, does not have this problem.

$$\frac{\left[\frac{H}{TR} - \frac{\overline{FA}}{\overline{TN}}\right] + 1}{2} \times 100$$

The score is calculated as follows:

1. the number of hits (real words the student marks as known (H)) is divided by the total number of real words on the test (TR);

2. the number of false alarms (pseudowords that the student marks as known (FA)) is divided by the total number of pseudo words on the test (TN);

3. the dividend in step 2 is subtracted from the dividend in step 1;

4. a constant of 1 is added to the result of step 3;

5. the sum of step 4 is divided by 2;

6. then the quotient in step 5 is multiplied by 100.

The result is a score on a scale of 0 to 100 which "estimates" the percentage of words from the list known by the student, but, like the regular Anderson/Freebody formula, it makes no claims about *which* words the student knows. However, as can be seen in the example in Table 2, the modified formula results in a more reasonable score for the student who circled all of the real words and some of the pseudowords.

[Insert Table 2 about here.]

Comparisons between the students' scores using the conventional Anderson/Freebody formula and the modified formula showed that the problem with the conventional formula--encountered when a student circled all the real words as known--occurred only with the on-grade level words. We correlated the students' scores using two scoring systems to see if they yielded different information. The correlations for on-grade level words (*r*s ranged from .52 to .87) were substantially lower than the correlations for the off-grade level words (*r*s ranged from .87 to .98). This indicated that the two scoring systems did yield different information for on-grade level words. In contrast, the scores for the off-grade level words were essentially the same, suggesting that the two scoring systems were assessing similar information. Because the modified formula corrects for the overestimation of students' vocabulary knowledge that occurred for the on-grade level words, the scores for the modified system were used in our analyses.

Procedures

The dependent measures were piloted in a third midwestern district. Students involved in the pilot testing were interviewed about the length and difficulty of the tests. The second grade students found the 60-item multiple choice test too long. Therefore, this measure was divided into two 30-item multiple choice tests and administered on subsequent days. Each fifth grade student, however, took one 60-item multiple choice test.

The tests were administered in classrooms by the regular teachers; they followed oral and written instructions supplied by the researchers. The second grade tests were administered in three sessions: (a) the yes/no test; (b) 30 items of the multiple choice test; and (c) the remaining 30 multiple choice items. The fifth grade tests were administered in two sessions: (a) the yes/no test; and (b) the 60-item multiple choice test.

The yes/no tests preceded the multiple choice tests so that the subjects would not gain information from the multiple choice tests which might affect their performance on the yes/no test. Since the yes/no format provides no choice other than "yes, I know it" or "no I don't," it is difficult, if not impossible, to "learn" vocabulary by completing it. Students were given as much time as they needed to complete the tests. The participating teachers reported that each testing session lasted approximately 20 minutes.

Results and Discussion

Overall Plan

The purpose of the present study was to investigate whether children in the second and fifth grades knew the meanings of vocabulary designated as "new" words by the teachers' manuals. The first set of analyses focused on students' performance on words from their current basal series by looking at their scores on the multiple choice tests and the yes/no tests. For historical comparisons, a second set of analyses was performed that involved using the results of the yes/no tests which compare student

performance on words from the series Gates used (Macmillan) with words from the series we used (Houghton Mifflin and Scott Foresman). Then, we compared scores from both the multiple choice and yes/no tests by converting them to a "common" scale. Finally, we analyzed vocabulary difficulty by comparing samples of "new" words from each series on an independent metric of difficulty, word frequency.

Multiple Choice Tests

Mean performance on the multiple choice tests (see Table 3) provided the primary data to answer the question of whether or not students knew the meaning of "new" vocabulary. The second graders' knew the meaning of more than 85% of the instructed second grade words. But, they also knew between 72% and 81% of the meanings of the "new," uninstructed, third and fourth grade words. The fifth graders knew the meanings of more than 78% of the instructed fifth grade words and more than 72% of the "new" sixth and seventh grade words. Our findings essentially replicate Gates' (1961, 1962) results. It is interesting to note that in spite of changes in basal reading materials over the last 25 years (Morris & Johns, 1987), students still know most of the "new" words. Our data indicate that the decreases in control over vocabulary in basal readers, which is equivalent to saying that more "new" words are introduced in the books, has not resulted in students knowing fewer of the "new" words.

[Insert Table 3 about here.]

As one would expect, the overall means of students' performance on the multiple choice tests decreased as the grade level of the words increased. Only the second grade students using the Scott Foresman series deviated from this pattern. In order to examine the overall trend, a one-way analysis of variance (ANOVA) with a repeated measures design was calculated for the words from each grade level for each reading series. The main effect for grade level of the words was significant in all cases: Houghton Mifflin Grade 2, F (2,27) = 29.05, MSe = 21.42, $p < .001$; Houghton Mifflin Grade 5, F (2,28) = 22.72, MSe = 15.45, $p < .001$; Scott Foresman Grade 2, F (2,45) = 99.61, MSe = 25.46, $p < .001$; Scott Foresman Grade 5, F (2,39) = 18.65, MSe = 21.54, $p < .001$. The Scheffe procedure was employed for post hoc comparisons (See Table 3 for means). All possible grade level comparisons of student performance on the multiple choice test for the Houghton Mifflin Series were statistically significant ($p < .05$). For the Scott Foresman second-grade students, performance on third- and fourth-grade sets of words, while differing significantly from the second-grade set ($p < .05$), did not differ from one another. The Scott Foresman fifth graders' performance revealed a different situation. There was no significant difference between fifth- and sixth-grade words but their scores on seventh-grade words were significantly different from their scores on the two lower grade level words. Even though the trends in the students' performance followed the expected pattern of decreasing as grade level increases, the changes are not as great as might be expected for words that are considered "new" to the students. Students were able to select appropriate meanings for a very high percentage of the "new" words.

In addition, correlations of the students scores on words from the three grade levels tested were calculated. The correlations were moderate to strong (rs ranged from .56 to .88), indicating that, students who performed well on words from one grade level, performed well on "new" words from the other two grade levels. Conversely, students who didn't know many of their on-grade level words also did not know many "new" words from subsequent grade levels.

Yes/No Tests

The second graders' mean performance on the yes/no test for second grade words indicated that they knew more than 85% of those words. Their performance on third- and fourth-grade words was somewhat lower; indicating that they knew about two-thirds of the upcoming "new" words. They scored between 63% and 71% correct on these "new" words. The fifth-grade students' performance on the yes/no tests showed that they knew more than 79% of the fifth-grade words. Their performance on

"new" sixth- and seventh-grade words showed they knew between 63% and 78% of these words (see Table 4).

[Insert Table 4 about here.]

The students' overall mean performance on the yes/no test decreased as the grade level of the words increased. This pattern was consistent for all four groups of students. In order to determine if this decrease in performance was significant, a one way analysis of variance using a repeated measures design was calculated for the students' performance on the words from each grade level for each series. The main effect for grade level was statistically significant in all cases: Houghton Mifflin Grade 2, F (2,27) = 263.69, MSe = 15.82, $p < .001$; Houghton Mifflin Grade 5, F (2,28) = 133.06, MSe = 17.19, $p < .001$; Scott Foresman Grade 2, F (2,44) = 272.67, MSe = 31.23, $p < .001$; Scott Foresman Grade 5, F (2,39) = 168.87, MSe = 16.89, $p < .001$. The follow up Scheffe procedure yielded mixed results regarding all possible grade level comparisons of student performance (see Table 4 for means). Second-grade students using the Houghton Mifflin series and fifth-grade students reading in the Scott Foresman series revealed a pattern of significant differences across grade levels ($p < .05$). The results were not as consistent for the other two groups of students. The second-grade students using Scott Foresman knew more second-grade than third- or fourth grade-words, but not more third- than fourth-grade words For the fifth-grade students using Houghton Mifflin, performance was similar for fifth- and sixth-grade words, but they knew significantly fewer seventh-grade words. As with the multiple choice tests, students' scores for on-grade level words correlated well with their scores on the words from higher grade levels (rs ranged from .60 to .80).

Yes/no tests were used in this study as a convergent dependent measure of students' vocabulary knowledge. The results of the yes/no tests did, in most cases, provide data that converged with the results of the multiple choice tests. For example, the students' overall mean performance on the yes/no tests mirrored the results from the multiple choice tests. On both measures, students knew more than 75% of the on-grade level words and more than 60% of the words from the two subsequent grade levels. Also, the results of correlations between the scores on the multiple choice test and the yes/no test show that students' performance on one test is related to their score on the other test. Correlation between the on-grade level scores from the multiple choice test and the corresponding yes/no tests were moderate (r = .39 -.64), and statistically significant ($p < .01$) for all but the fifth-grade students using the Houghton Mifflin Series (r = .31 $p > .05$). The correlations between students' multiple choice scores and yes/no scores for "new" words from subsequent grade levels were moderate to strong (r = .44 - .73), and all these correlations were statistically significant ($p < .01$).

Historical Comparisons.

In order to make comparisons in student performance on words from the series Gates used (Macmillan) and words from the series we used (Houghton Mifflin and Scott Foresman), students were tested with a yes/no test on a random sample of words from the Macmillan (1958) basal series. The students' performance on the words from the Macmillan series were compared to their performance on words from the corresponding grade levels in their own series using dependent sample t-tests (See Table 5). The overall mean score for second graders is higher on Macmillan words than it is for the words from their own series. But, the difference is statistically significant only in the Scott Foresman Series. The performance of fifth graders is reversed, their overall mean score is higher for words from their own series, whether it was Scott Foresman or Houghton Mifflin. These students indicate knowing more of the "new" words from their own basal reading series.

[Insert Table 5 about here.]

An interpretation of these data is that, while the means for the current basal words differ *statistically* from the Macmillan means, the differences do not represent an *educationally* significant difference. To put it another way, while the fifth grade means for the current basals (Houghton Mifflin M = 78.74;

Scott Foresman M = 80.78) were significantly higher than those for the Macmillan words (M = 73.59 and M = 74.78 respectively), in reality, the difference between the means reflects a child getting only one more word right. The same is true for the second grade scores (See Table 4 for means).

Analyses for Both Multiple Choice and Yes/No Tests

Obviously, the two metrics--multiple choice and yes/no--yielded somewhat different estimates of overall vocabulary knowledge, with the yes/no test yielding generally lower estimates. To bring the metrics into better "balance," they were converted to a "common" scale by dividing each above-grade level proportion correct by the on-grade level proportion correct based upon the reasoning that the on-grade level proportion represents an "effective" ceiling on performance. These results (presented descriptively in Table 6) further illustrate that, although student performance decreased as grade level increased, the students still knew relatively high proportions of the words irrespective of test format.

[Insert Table 6 about here.]

Word Frequency Analyses

The data presented so far have a mixed appearance. On the one hand, differences from one grade level to the next were, by and large, statistically significant. On the other hand, the data, showed that students could select appropriate meanings for a very high percentage of words as much as two levels beyond their current grade level. Thus, how much harder are words at higher grade levels?

In order to determine the level of difficulty of the "new" words and whether it affected students' performance on a test of vocabulary knowledge, the various samples of words were compared on an independent metric of difficulty--word frequency (Carroll, Davies, & Richman, 1971). Word frequency has a long standing tradition as a predictor of vocabulary and comprehension difficulty, with the obvious and consistent result that the less frequently a word occurs, the less likely that its meaning will be known. These analyses were completed using SFI values which are logarithmic transformations of actual frequency values (U values).

Students' scores on each item were compared with the SFI values for the words in order to determine the degree to which the frequency of a word predicts student performance. The correlation between frequency and performance was statistically significant ($r = .38$, $p < .001$), but only explained about 14% of the variance in test performance.

A 3 (Series) x 2 (Test Level--primary or intermediate) analysis of variance (with separate tests of words by Grade Level nested within Test Level) of SFI values was calculated for Series by Test Level (See Table 7). Words from the Macmillan (1958) series, that Gates used in his studies, were included as the third series in order to look at whether words have actually become more difficult in the last 25 years. The Test Level variable was an aggregated mean of SFI values across several grade levels of words. Test Level 2 consisted of words from grades 2, 3, and 4 for all series. Test Level 5 consisted only of words from grades 5 and 6 because the Macmillan (1958) series did not include seventh grade words. There was a significant main effect for Test Level ($F(1) = 26.27$, $p < .001$), indicating that the words given to our fifth graders were "harder" (i.e., less frequent) than those given to second graders. There were no grade level differences (e.g., 2 versus 3 versus 4 or 5 versus 6) within either test level. The main effect for Series and the interaction of Test Level by Series were not significant. What these results suggest is that as grade level increases, words do become less frequent, but at best, the change is modest from one grade level to the next. Also, the lack of a series effect suggests that vocabulary has not become more or less difficult in the last 25 years.

[Insert Table 7 about here.]

The results of this study are meaningful only if it is true that a representative sample of words was chosen. In order to determine if the sample chosen for this study was a representative sample from the basal lists of "new" words, a second random sample of words was chosen for each series, and the SFI values for the two samples from each series were compared by calculating t-tests. The results showed no significant difference between the two groups of words for any of the series. This strengthens the claim that the words used in the present study are a representative sample of the words on the lists of "new" vocabulary.

Conclusion

The data from both the multiple choice tests and the yes/no tests indicate that children *did* know many of the "new" vocabulary words in their basal reading series. The results showed that, on average, children knew more than 70% of the "new" words that were sampled from their basal reading series. While there is no basis for claiming that these children had complete knowledge of these words, the children's scores on these tests indicate that they had at least partial knowledge of many of the words.

Although student performance decreased as the grade level of the words increased, and this difference was statistically significant, this finding must be viewed in terms of "real world" differences. In most cases, the difference in performance from one grade level to the next was only one or two words, which is not necessarily a big difference considering that these students knew such a high proportion of all the words tested.

The similarity between the number of "new" words children knew in the studies Gates conducted in the late 1950's and early 1960's and the number of words children knew in this study raises questions about where children learn these words and whether or not the words targeted for instruction in basal series have changed much in the last 25 years. Answers have been offered by researchers regarding where and how children learn new words. The results from this study substantiate the view that children learn words through encountering them in various contexts (Nagy, Anderson & Herman, 1987), such as independent reading, conversations, the media, and school reading in the content areas.

Although the premise of this study is that children know the meaning of many of the new words in their basal series, it is necessary to remember that the words selected by basal publishers for instruction prior to a reading selection may be included for reasons other than the teaching of word meanings. In response to our request for information, an executive editor from one basal company explained which words were included in lists of "new" vocabulary: key words for understanding the passage, words that could be used to apply recently learned skills, and words that are not likely to be in the students' oral vocabularies. A teacher who is aware of these alternative purposes of vocabulary instruction will not waste valuable lesson time on teaching meanings the students already know. To avoid cases where teachers are not aware of these purposes, publishers should explain in the teachers' manuals why each word is included in a "new" vocabulary list. This, in combination with assessment, will allow teachers to maximize instruction and meet individual needs.

The results of this study suggest topics for further research. First, research on measures for assessing vocabulary knowledge might help determine what sorts of word knowledge the yes/no and multiple choice tests are measuring and thereby provide information about the uses for each type of test. Second, research on vocabulary instruction might explain why students receive similar test scores for words they have been taught and words they have not been taught in school.

Finally, a cautionary note for researchers is necessary. It was found that the Anderson/Freebody formula used for calculating scores on the yes/no vocabulary test is inappropriate under certain conditions. It fails to discriminate between students who indicate that they know all of the real words irrespective of the number of pseudowords that they also say they "know." By modifying the formula in a manner more consistent with traditional correction for guessing procedures on true-false items, we

found that correlations between yes/no tests and multiple choice tests improved. Researchers are advised to take these findings into consideration when they use the Anderson/Freebody yes/no test.

In conclusion, children did not demonstrate a great deal more knowledge for the words they have been taught in school than for words they had not yet been taught. This suggests that direct instruction can account for only a small portion of the new words children are learning. Basal publishers and teachers need to be cautious about making assumptions regarding which words children know or do not know at each grade level. And they may wish to differentiate vocabulary instruction based upon the different reasons for including words in a "new" vocabulary list.

References

Aaron, I. E., Jackson, D., Riggs, C., Smith, R. G., & Tierney, R. (1983) *Scott, Foresman Reading.* Glenview, IL: Scott, Foresman and Company.

Adams, M. J., & Huggins, A. W. F. (1985). The growth of children's sight vocabulary: a quick test with educational and theoretical implications. *Reading Research Quarterly, 20*(3), 262-281.

Anderson, R. C., & Freebody, P. (1981). Vocabulary Knowledge. In John T. Guthrie (Ed.) *Comprehension and Teaching: Research Reviews,* (pp. 77-117), Newark, DE: International Reading Association.

Anderson, R. C., & Freebody, P. (1982). *Reading comprehension and the assessment and acquisition of word knowledge* (Tech. Rep. No. 249). Urbana: University of Illinois, Center for the Study of Reading.

Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Hutson (Ed.) *Advances in reading/language research,* (pp. 231-256), Greenwich, CT: JAI Press.

Anderson, R. C., & Freebody, P. (1985). Vocabulary knowledge. In H. Singer & R. Ruddell (Eds.), *Theoretical models and processes of reading* (3rd Ed, pp. 343-371). Newark, DE: International Reading Association.

Anderson, R. C., & Kulhavy, R. W. (1972). Learning concepts from definitions. *American Educational Research Journal, 9* (3), 385- 390.

Beck, I. L., & McKeown, M. G. (in press). Vocabulary learning and instruction. In P.D. Pearson (Ed.), *Handbook of reading research, Vol. 2.* New York: Longman.

Carroll, J. B., Davies, P., & Richman, B. (1971). *American Heritage Word Frequency Book.* Boston, MA: Houghton-Mifflin.

Duffelmeyer, F. A. (1984). The effect of context on ascertaining word meaning. *Reading World,* Oct., 103-107.

Durkin, D. (1984). Is there a match between what elementary teachers do and what basal reader manuals recommend? *The Reading Teacher, 37,* 734-744.

Durr, W. K., LePere, J. M., Pikulski, J. J., & Bunyan, R. P. (1983). *Houghton Mifflin Reading Program.* Boston, MA: Houghton Mifflin Co.

Gartler, M., & Benditt, M. (1958). *Macmillan Reading Series.* New York, NY: Macmillan.

Gates, A. I. (1961). Vocabulary control in basal reading material. *The Reading Teacher, 15* (2), 81-84.

Gates, A. I. (1962). The word recognition ability and the reading vocabulary of second and third-grade children. *The Reading Teacher, 15* (6), 443-448.

Gipe, J. (1978-79). Investigating techniques for teaching word meanings. *Reading Research Quarterly, 14,* 625-644.

Johnson, D. D., & Pearson, P. D. (1984). *Teaching Reading Vocabulary.* New York: Holt, Rinehart and Winston.

Karlsen, B. (1980). *There is no such thing as a third grade word.* Paper presented at the California Reading Association Annual Convention.

Kolstad, R. K., Briggs, L.D., Bryant, B. B., & Kolstad, R. A. (1983). Complex multiple-choice items fail to measure achievement. *Journal of Research and Development in Education, 17* (1), 7-11.

McKeown, M. G. (1985). The acquisition of word meaning from context by children of high and low ability. *Reading Research Quarterly, 20* (4), 482-496.

Morris, J. A., & Johns, J. L. (1987). Are 1st grade reading books easier than 20 years ago? *The Reading Teacher, 40* (4), 486-487.

Nagy, W., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.

Nagy, W., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*, (2), 237-270.

Nimmer, D. N. (1983). Multiple true-false classroom tests. *Clearinghouse, 56* (6), 257-258.

Searls, E., & Klesius, J.P. (1984). Multiple meaning words for primary children and ways to teach them. *Reading Psychology: An International Quarterly, 5*, 55-63.

Author Notes

We gratefully acknowledge William F. Nagy for showing us how to use his computer program to create multiple forms and score the yes/no vocabulary tests. Also, we would like to thank Jeanette Allison Hartman and Paul Kerr for helping us with the tedious task of data entry.

Table 1

Sample Multiple Choice Items

Grade-Item Type	Item	Distractor Type[a]
2nd-Example	Which of these is a tune?	
	A. A radio	1
	B. A song	*
	C. A shiny stone	4
	D. A fish	3
	E. A loud noise	2
2nd-Attribute	How might you feel if you were furious?	
	A. Hairy	3
	B. Tired	1
	C. Angry	*
	S. Upset	*
	E. Smart	4
5th-Definition with context	The sly cat ate the hamburger while the man was gone from the table.	
	What does sly mean in this sentence?	
	A. Clever	*
	B. Tricky	*
	C. Hungry	2
	D. Slow	1
	E. Fat	4
5th-Definition no context	Communicate means	
	A. go to jail.	4
	B. travel a long distance to work.	3
	C. deliver some equipment.	2
	D. give or exchange news.	*
	E. talk or share information.	*

[a]Distractor Types
*keyed response
1. same semantic category as target word
2. same syntactic category as target word
3. graphically similar to a decoding miscue of the target word
4. unrelated to the target word

Table 2

Comparisons of Student Scores on the Yes/No Test Using the Anderson/Freebody Formula and the Modified Formula

Real Words Circled	Pseudowords Circled	A/F Formula	Mod. Formula
90/90	0/90	100%	100%
90/90	30/90	100%	83.34%

Table 3

Mean Vocabulary Scores on the Multiple Choice Tests

Basal Series	Word Grade Level	M	SD
Houghton Mifflin	2	86.68	8.06
Second Grade (n = 28)	3	81.93	8.99
	4	77.25	8.35
	Total	81.95	7.59
Houghton Mifflin	5	82.00	5.11
Fifth Grade (n = 29)	6	79.07	6.19
	7	75.07	5.62
	Total	78.71	4.66
Scott Foresman	2	86.62	6.19
Second Grade (n = 45)	3	72.69	8.39
	4	74.76	9.48
	Total	78.02	7.02
Scott Foresman	5	78.78	9.66
Fifth Grade (n = 40)	6	78.03	9.92
	7	72.95	10.82
	Total	76.58	9.41

Table 4

Mean Vocabulary Scores on the Yes/No Tests

Basal Series	Grade Level Words	M	SD
Houghton Mifflin	2	89.71	5.98
(*n* = 28)	3	71.00	6.87
	4	66.79	6.95
	2-4	75.83	5.77
Macmillan	2-4	76.50	7.59
Houghton Mifflin	5	79.86	6.82
(*n* = 29)	6	77.62	9.60
	7	63.48	5.74
	5-7	73.66	6.76
	5,6	78.74	7.61
Macmillan	5,6	73.59	5.28
Scott Foresman	2	87.91	7.50
(*n* = 45)	3	64.31	10.47
	4	63.87	11.15
	2-4	72.03	8.71
Macmillan	2-4	76.76	9.78
Scott Foresman	5	82.80	7.17
(*n* = 40)	6	78.75	6.55
	7	66.58	9.34
	5-7	76.04	7.02
	5,6	80.78	6.51
Macmillan	5,6	74.78	5.78

Table 5

Dependent Sample T-tests for Words from Macmillan and the Students' Own Reading Series

Group	n	SS	t	p
Grade 5, 6				
Houghton Mifflin	29	664.55	5.70	.001
Scott Foresman	40	646.00	9.32	.001
Grade 2, 3, 4				
Houghton Mifflin	28	603.11	-0.75	.462
Scott Foresman	45	1173.18	-6.14	.001

Table 6

Means/Vocabulary Scores Converted to Proportions of On-Grade Level Scores

Basal Series	Grade Level Words	Yes/No	Multiple Choice
Houghton Mifflin	2	100.00	100.00
	3	79.14	94.52
	4	74.45	89.12
	5	100.00	100.00
	6	97.20	96.43
	7	79.49	91.55
Scott Foresman	2	100.00	100.00
	3	73.15	83.92
	4	72.65	86.31
	5	100.00	100.00
	6	95.11	99.05
	7	80.41	92.59

Table 7

Analysis of Variance for Standardized Frequency Index by Test Level

Source	df	SS	MS	F	p
Level	1	2,339.95	2,339.95	26.27	.001
Series	2	66.69	33.35	0.37	.69
Level*Series	2	145.88	72.94	0.82	.44
Error	241	21,463.61	89.06		
Total	246	24,148.18			